

Integrating New Technologies: Electronic Distribution

What are the strategies being proposed by suppliers of Electronic Distribution product for dealing with paper and electronic delivery requirements?

by Frank Gilbane

Welcome to the first installment in our new a series on integrating new technologies. The focus of the new series will be on issues that information system managers who are implementing CALS or CALS-like solutions today have to grapple with in applying and integrating new technology solutions. We will invite authors with different points of view to write short articles that describe their approach for dealing with a particular technical issue. Our goal is to help you consider the critical issues when making decisions on integrating new technology solutions. The topic covered in this issue is electronic distribution.

A goal of virtually all CALS or CALS-like strategies is to be able to manage information that has traditionally resided in paper documents as effectively as information typically managed in corporate databases, and to distribute the information in either paper or electronic form. One of the thornier issues to resolve is how to store the information in a database so that it can be retrieved and presented in suitable form. Directly related to this issue is the question of what constraints this approach imposes on the format of the information. For example, there is a cost involved in supporting paper and electronic

"pages" of information that not only contain the same content, but look the same. Is this a necessary or acceptable cost? Is it a user requirement?

We have asked a number of technology suppliers to present their views on this particular aspect of an electronic distribution strategy. These are critical issues for any electronic distribution strategy and have a major affect on the ability of the strategy to meet user requirements, meet financial constraints, and protect the information from obsolescence.

Notice that we are not covering IETMs (Interactive Electronic Technical Manuals) specifically here—electronic distribution and the question relating to extracting and formatting information for presentation is only one of the areas a full blown IETM implementation must deal with.

To help you evaluate the three views presented in the context in which they were written, we have provided a quote from our request to the authors. We asked them to assume they were presenting their approach to an IS management group, this...

...IS team is attempting decide whether to advise adoption of an approach that either: ensures paper and electronic pages are identical or similar, requires that electronic information be formatted differently for different media, or provides a hybrid solution. They are confused about how SGML fits into these approaches and how important it may or may not be. What would you say to them?

You will not be surprised to find different responses from different vendors, and there are still other views on the issue from vendors not represented here. We believe that the selected articles provide you with an overview of the spectrum of choices. Frank Boosman presents a case for a solution based on new file format designed for PostScript™-based viewing, and Tim Bray and Moira Meehan's discussions of the issues suggest a more hybrid approach. You may find it interesting that SGML plays a role in all of the approaches, although the authors differ in exactly what that role is.

Remember that the right approach for you will

Frank Gilbane

is president of Publishing Technology Management, Inc. He consults to government and industry on the application of publishing, information management, and standards technology. He has been an active member of the CALS ISG since 1987 and has been involved with applying computer technology to business problems since 1974. Frank is also a member of the CALS JOURNAL Editorial Advisory Board.



depend on your own user requirements, and that, while most of us are concerned with CALS-like requirements, there are other, less demanding needs for long term information management. For example, distributing a novel for viewing on a portable CD-ROM is a very different problem than distributing dynamic technical and logistics information residing in distributed databases. The same careful analysis of your needs and the technical questions is required, but the answers you come up with may be different.

We encourage you to respond both to *CALS JOURNAL* and to the authors directly with any questions or reactions you have.

Electronic Document Communication: A Cross-Platform Approach

by Frank Boosman

As an evolving standard, CALS and the tools to implement it are not yet complete. We at Adobe Systems are developing products that we believe can help make CALS work for millions of desktop computer users on a daily basis. Adobe Systems is developing a new family of products called Adobe™ Acrobat™ that makes effective universal document transmission—or *document communication*—possible for the first time. Adobe Acrobat ensures that documents communicated via any electronic medium retain the integrity of their originals. The result is that users can freely transmit and receive documents without regard to computer hardware, operating system, or software application used.

Below is a description of the components of these products and how they work together to overcome the major barriers to document communication.

Frank Boosman

is a Senior Product Marketing Manager of Adobe Systems Incorporated. He is responsible for new products and technologies spanning personal computer and workstation platforms in the field of document communication. Before the acquisition of Emerald City Software by Adobe, he was the Director of Product Planning and managed the Smart Art tool for the Macintosh.

ADOBE ACROBAT COMPONENTS

Adobe Acrobat is a family of products that work together to enable document communication. The key individual components include:

Acrobat Viewer—a software program that enables users to view, navigate, annotate, and print documents represented in the Portable Document Format (PDF), Adobe's open, PostScript-based file format for cross-platform electronic documents. Acrobat Viewer software will be available for Macintosh, Windows, DOS, and UNIX platforms.

Acrobat Distiller™—a software program that translates PostScript files into Portable Document Format files. Acrobat Distiller will be available for Macintosh, Windows, and UNIX platforms.

Acrobat PDF Writer—a specially designed printer driver that produces PDF files directly from applications. PDF Writers will be included with Acrobat Viewers for Macintosh and Windows.

THE PORTABLE DOCUMENT FORMAT

The key to Adobe's approach to document communication is the Portable Document Format. A PDF file represents the final form of a document. It describes documents containing any combination of text, graphics, and images.

Like the PostScript language, PDF files are both device- and resolution-independent. This means they can represent information about a document's appearance that can be rendered by all major computer display or output devices. This is essential in today's world of multi-platform computing environments.

The PDF format, like the PostScript language, adheres to the seven-bit ASCII (American Standard Code for Information Interchange) standard recognized by all computer hardware and software makers as a universal form of electronic information interchange.

CREATING PDF FILES

Because the Portable Document Format is hardware-, operating system-, and application-neutral, virtually any document on any platform from any application can be converted to a PDF file.

A PDF file can be created from any software application that supports PostScript printing, such as word processors, graphics programs, databases, or spreadsheets. It is converted into a PDF file by using the Acrobat PDF Writer printer driver or Acrobat Distiller.

SMALL, PORTABLE FILES

The PDF file format is seven-bit ASCII, which makes it extremely portable between diverse hardware and operating system environments. While seven-bit ASCII formatting ensures universal transmission of PDF files, data

compression results in fast transfer and low consumption of disk and memory resources.

PDF files can be communicated using virtually any existing or future storage or transmission medium. The only universal representation previously available for this type of cross-platform communication was ASCII text. PDF provides the first universal alternative to ASCII as a final form document representation.

VIEWING DOCUMENTS

Acrobat Viewer software is used to view, navigate, annotate, and print PDF files. Its graphical interface provides clear, intuitive access to all of its functions. The applications used to create the PDF files are not required to view them.

The Acrobat Viewer includes navigational tools that can scroll, zoom, and allow the user to access different portions of a document using miniature "thumbnail" representations of each page in the document. The user can leaf through the thumbnails like through a magazine, then jump from the page on-screen to a distant page by mouse-clicking the thumbnail for the destination page.

The Acrobat Viewer also has "live links" features that enable users to create hypertext-like links in PDF files. These are ideal for cross-references within a document.

Additionally, Acrobat Viewer software allows the reader to annotate PDF files by affixing electronic text notes to them. These notes are transmitted along with the main PDF file and may be hidden or displayed at the reader's discretion.

FONT SUBSTITUTION

A key component of the Adobe Acrobat software is a font substitution capability that solves a fundamental barrier to document communication, the "font problem." In current systems, if a computer lacks a font in a document received, the user may not be able to view or print the document, or the computer may substitute a different font for the absent one, drastically altering the document's appearance or making it completely illegible. When the receiver of a PDF file lacks a font used in a document, Acrobat mimics the absent fonts with a special Adobe Type 1 font. The appearance of the unavailable font is approximated so closely that the look and feel of the document is completely preserved.

ACROBAT AND THE CALS INITIATIVE

For the CALS community, Acrobat offers compliance with important existing standards. First, its file format—the Portable Document Format—will be an openly published standard. This format is based on PostScript, a published standard supported by dozens of hardware manufacturers and thousands of software applications.

Second, the Standard Page Description Language (SPDL), the preferred final-form representation for CALS documents, is based on PostScript. Since the various components of Acrobat are PostScript-based, this will make translation and migration between SPDL, PostScript, and PDF as simple as possible.

Third, the font technology that is a key component of both PostScript and Acrobat, the Type 1 font format, has been sanctioned by the International Organization for Standardization (ISO) as the standard format for electronic fonts. Type 1 is also a part of the CALS specification.

Finally, Adobe recognizes the important role SGML has to play in the world of document communication. SGML excels at the representation of structure of documents but lacks the ability to precisely represent and reproduce the appearance and format of documents. Acrobat provides just such a capability, and to make PDF and SGML synergistic, Adobe is building software that will allow SGML users to combine their structural data with PDF files to create even more useful electronic documents.

Acrobat will bring immediate benefits to CALS users. For example, the first government organization projected to adopt Acrobat, the Defense Printing Service (DPS), will store its PostScript documents as PDF files. This will give them the flexibility to permit Air Force customers with ATOS data, Army customers with JCALS data, and Navy customers with ADMAPS data to access and view their documents from any desktop computer, then send them to DPS for re-publishing in whole or in part, saving time and money.

In short, Acrobat offers the CALS community both compliance with standards and compatibility with existing desktop computers and software applications. This, we believe, will give Acrobat an important role in advancing the CALS initiative and bringing its benefits to users' desktops. ■

SGML, Open Systems, and Flexibility

by Tim Bray

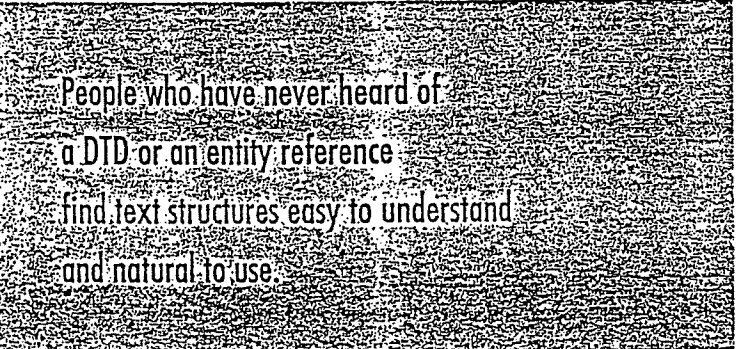
We use computers for the sake of speed and flexibility. Flexibility is the more important of the two because it makes information re-usable. Information is expensive to create—too expensive to use just once.

I feel that the biggest benefits of the CALS program will come from its emphasis on flexibility. This is why



SGML is so important: a graphic is just a graphic, and a drawing just a drawing, but with SGML, textual information can be many things to many people.

Open Systems is another technology trend that is all about flexibility. Open Systems thinking is based on the premise that no single piece of technology can solve all the problems; tools and applications must be combined to get the job done. Open Systems make it possible to choose the unique set of tools that best solves the problem and provides the flexibility to build them into a solution.



People who have never heard of
a DTD or an entity reference
find text structures easy to understand
and natural to use.

In this article, I will show how SGML and Open Systems-based tools, working together, can deliver new levels of performance to users of electronic information, especially information of the type found in the CALS community.

FIRST, YOU HAVE TO FIND IT

One thing is certain about CALS information: there's a lot of it. Also, the amount is growing at a frightening rate. This means that any electronic dissemination system must have very powerful full-text retrieval capabilities and not be frightened of a few gigabytes here or there.

This information is complex, containing a variety of graphical and structured-text information. In real-world texts, structures are complicated, hierarchical, and big; SGML supports this well. It is important to realize that in this environment, the very notion of a "document" is weakened. A document can be a 5-line compliance memorandum, a 40,000 page maintenance volume set,

Tim Bray

is Senior Vice President—Technology, and a founder, of Open Text Corporation. He has worked for Digital Equipment Corporation, GTE, and was manager of the New Oxford English Dictionary Project at the University of Waterloo in Canada. This project, concerned with the problems of making a highly-structured, 100-year-old, gigabyte-scale text database useful to a highly nontechnical user community, led to the invention of the technology now marketed by Open Text Corporation.

or anything—literally *anything*—in between.

Conventional full-text systems provide *document retrieval*; they assume that information comes neatly divided into reasonable-sized chunks, and that people will be satisfied if they can retrieve these.

This is a real pity because our experience has shown that SGML is a wonderful framework for information retrieval. SGML assumes a universe of textual and graphical material populated with many structural elements; the idea of dividing this up into "documents" or "records" is fundamentally wrong. It is much better to control searches using any or all of the structural elements and to retrieve information by whichever unit is most convenient.

We have found that ordinary users, people who have never heard of a DTD or an entity reference, find text structures easy to understand and natural to use. Figure 1 shows Open Text's PAT text database system at work. The user has built a complex query involving string-matching in three levels of structure to find a small set of information in a large database. This query would be impossibly difficult using standard boolean document retrieval, but the clean structure provided by SGML makes it easy. What the figure doesn't show is that the whole sequence took just a few seconds.

Then, You Have to Look at It

Finding the information you need is tough without the right tools. Using it can be even tougher. The material is complex, there's a lot of it, and it can contain graphics, formulas, and other material that is difficult to display.

Luckily, SGML provides the solution to half this problem, and Open Systems take care of the other half. SGML's clean, consistent text markup allows the construction of text viewers that can provide many different presentations of the same text, computed (fast!) in real-time. Figure 2 shows Open Text's LECTOR display tool providing several views of the same information, optimized for different purposes.

But an SGML viewer can't solve all the problems of display, particularly in a typical CALS environment with heterogeneous text and graphics formats. Although the structured-text community tends to look down on "WYSIWYG," there are times when a document must be presented *just right* to be useful.

This is where Open Systems comes in. No one tool can solve all our display problems. But that's OK, because with Open Systems we don't have to use just one tool. For example, Open Text Corporation doesn't market an image display system, but most of our customers use one. Figure 3 shows an example. Here the PAT text database finds information and uses another tool (In this case Precise/Review, from Precise Systems) to provide high-performance image display.

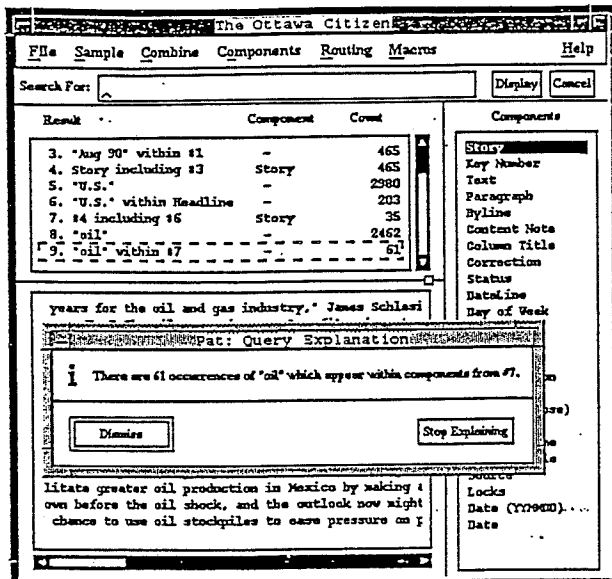


Figure 1. How many "Story" structures containing an August 1990 date-stamp had "U.S." in the headline and mentioned "oil" in the text? SGML and fast full-text search make them easy to find.

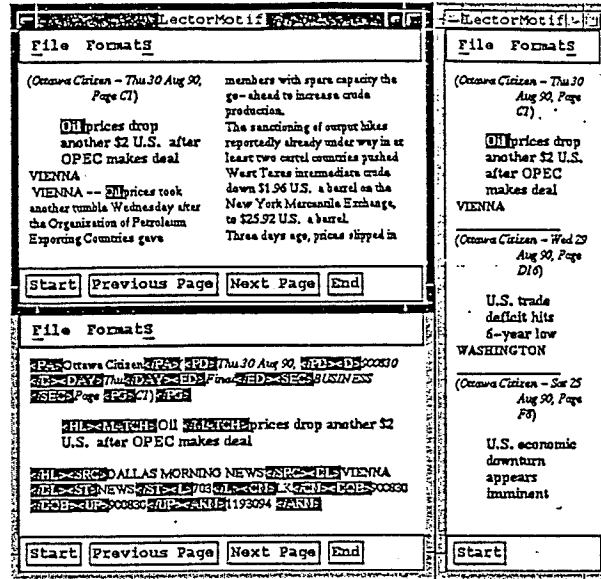




Figure 2. Some of the results from the query in Figure 1, presented three different ways to meet the needs of three different jobs.


SHAPING THE GLOBAL LANGUAGE




The InterCAP Publishing Suite




**A Leading Edge
Drafting Tool**



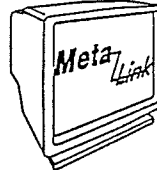
**A Dedicated Technical*
Illustration System**



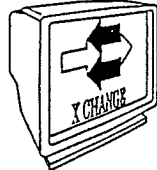
**A Robust View And
Markup Package**



**An Illustration Package*
For Quick Revisions**



**A CGM-Based Graphics
Authoring System**

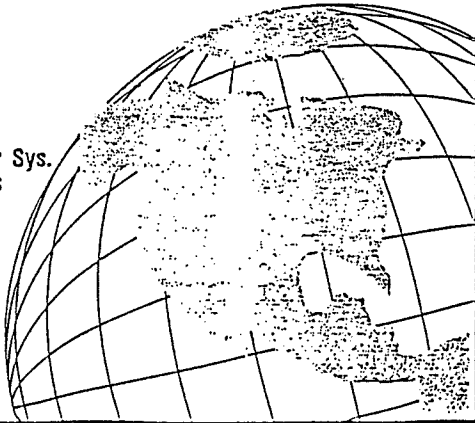


**An Easy-To-Use
Data Converter**

Partial Customer List:

Airbus Industries (All Partners) Allison Gas Turbine AT&T Bendix Oceanics British Aerospace Computer Sciences Corp. Caterpillar Daimler Benz John Deere E.I. Du Pont De Nemours & Co. Fiat	General Electric Grumman Hyster Hewlett-Packard Lockheed LTV Massey-Ferguson Motorola NCR Opel Raytheon	Rockwell Rohr SNECMA Sundstrand Power Sys. Texas Instruments Unisys U.S. Army U.S. Navy U.S. Air Force Westinghouse Xerox
--	---	---

* Selected as Baseline Products for CSC/JCALS



InterCAP Graphics Systems ♦ 116 Defense Highway ♦ Annapolis, MD 21401 ♦ 410.224.2926 ♦ FAX 301.261.8358

Principal Trust

File Sample Combine Components Routing

Search For:

Result	Component
1. "when the work load red"	-
2. "when the"	-
3. "when"	-

These amounts when added to the losse
These amou
so far is t
hould hav
ould hav
nsation pla
nsation pla
time to ti
time to ti

Principal Trust Litig

File Formats

000001.2000001
000002

Hon. Connie Osterman

It is very important that inspectors examining the affairs of our trust companies be cognizant of the Act and its interpretations. It is also important to be aware of changes occurring from month to month in a trust company's operation. This can only be accomplished if inspectors are employed by the Trust Program. From time to time, when the work load reduces, an inspector could be loaned to assist some other program which requires additional assistance.

With the number of requests for information and the complaints received during the year, it is imperative that a Loan Companies Act be implemented without delay. Alberta is the only major province without such legislation to effectively administer and control companies carrying out business in this province as mortgage companies and mortgage brokers.

During the year, 3 extra-provincial trust companies were registered in this province; also, 2 extra-provincial trust companies amalgamated, for a total increase in registrations of 2. The total number of trust companies now registered in the province is 53, less 2 Alberta-incorporated companies not carrying on business with the public.

A review of the enclosed reports indicates that there is nothing of a serious nature to report concerning our 7 Alberta-incorporated trust companies. Minor infractions will be reported on separately in my individual reports on each company.

Generally speaking, with the exception of [redacted] and Principal Savings and Trust Company, our Alberta-incorporated trust companies have fared better than their counterparts across Canada. It will be anyone's guess what the situation will bring in 1982 should the erratic financial market conditions prevail, as well as the unprecedented high interest rates which have affected all financial institutions operating in Canada.

[Handwritten Signature]

T. E. Dansereau

Encls.

cc: J. Barry Martin
T. Salob

Start Previous Page Next

Image: 000001.2

Review: principal

000002

Hon. Connie Osterman -2- November 30/82

Doc 1 of 1 Page 1 of 1

Figure 3. This database is full of newspaper clippings, handwritten notes, and the like; the solution requires smooth integration between text database and imaging technology.

WE HAVE THE TECHNOLOGY

Initiatives such as CALS are placing a lot of information on-line. This doesn't automatically mean that it is available or useful. CALS, however, because of its general emphasis on flexibility and Open Systems, and specific reliance on SGML, is an environment where on-line information can be a dynamic, productive, vital resource. And the technology to achieve this is here today. ■

Interleaf

by Moira Meehan

As organizations increasingly move to on-line distribution to cut costs and improve the quality of their information, they have the opportunity to re-engineer

the information in order to maximize its value. They can choose to distribute information in electronic versions that exactly replicate the original hardcopy, pay no attention to the original, or steer a middle course that replicates the "look and feel" of paper documentation but does so in a way optimized for electronic display. As they make this decision, they are typically choosing from three basic data types that underlay the vast majority of distribution products: raster images, a page description language, or SGML. These three different distribution approaches have their own strengths and weaknesses—including costs—and therefore each is appropriate in some instances and not in others. The key is to understand the implications of each choice before making the decisions.

Moira Meehan
is a SGML/CALS Product Marketing Manager at Interleaf, Inc.

ELECTRONIC AND INEXPENSIVE

"I don't want any paper on my plane except the flight checklist. I don't care if you turn every page into a raster image, as long as the pilots aren't carrying the weight of the paper." This was the position of an Air Force Program Manager speaking for a System Program Office recently. He went on to say that the lowest cost digital solution would be accepted since funding was tight, even if it meant the data would not be delivered in CALS formats.

Creating a simple raster, perhaps in CCITT, is by far the least expensive mode for creating electronic files

Where hardcopy matching is not a requirement,
SGML-based systems can give tremendous flexibility
and software independence.

from existing hardcopy documents. This approach would certainly solve the problem weight the program officer is concerned with, and the electronic version is guaranteed to look exactly like the original.

This solution has its drawbacks, however, because it is truly the equivalent of creating a microfiche image with no true "intelligence" for helping the user locate information—the user needs to know what information she needs and where to find it. There is no ability for full text searches or hyperlinking between sections. While it is possible to embed key words into the resulting file for some searching, this is necessarily manual and is therefore error prone. Also, because the image is a raster, it may lose definition when zoomed and therefore be difficult to read.

So, while this approach may be the least expensive, it is also the least effective. This solution is most useful where the documents exist in hardcopy and are unlikely to undergo significant revision.

ELECTRONIC AND INTELLIGENT

If low cost electronic pages are the goal, a more intelligent and readable file can be created from a page description language (PDL). PDLs are used as the intermediate format between the text processing tool and the laser printer or typesetter. Unlike a raster image, a PDL file carries with it all the document's text (and not just a raster image of the text). This means that a full text retrieval engine, such as Verity or Fulcrum, can enable

1388

Concurrent Engineering from Design to Implementation

PSI's family of R&M software pulls data from CAE/CAD or database programs, performs R&M analyses, prints reports, and then sends your predictions to any LSA software package. We cover your R&M needs from start to finish.

- Commercial and Military Reliability Predictions
- Maintainability Prediction
- Failure Mode, Effects and Criticality Analyses
- Output to LSA

CIRCLE READER SERVICE NUMBER 93

Put us to the test. Call for a FREE DEMO!

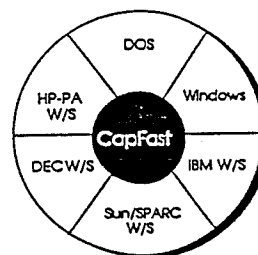
POWERTRONIC SYSTEMS, INC.

13700 Chef Menteur Hwy. New Orleans, LA 70129

504-254-0383 FAX 504-254-0393

CapFast™

Starts from \$395.00



World-class Schematic Design and Interface Tools

Phase Three Logic's CapFast family of EDA tools have been used by many leading companies such as Delco/GM, Analog, Supercomputers, Sharp, Allen Bradley and many more major companies world-wide. CapFast includes features such as an advanced hierarchical schematic editor that allows both top-down and bottom-up design approaches, an on-line electrical rules checker, an interactive simulation grapher, extensive parts library, intelligent packager, symbol creation editor and interfaces to SPICE, Verilog, HILO, Actel and Xilinx as well as to popular PCB design packages, like TANGO, PADS, PCAD. Schematic and symbol translator options based on the EDIF 200 standard available at nominal prices.

☎ Call Sales/Marketing (503) 645-0313 or Fax (503) 645-0207



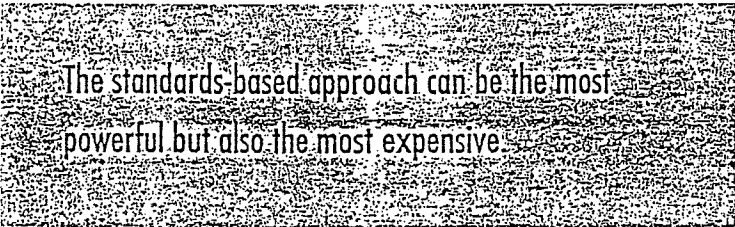
Phase Three Logic, Inc.
1600 NW 167th Place
Beaverton, Oregon 97006 - U.S.A.

CIRCLE READER SERVICE NUMBER 63



users to quickly find any set of words in what may be a huge set of PDL files. And because the PDL also continues the font metrics, files can be easily scaled and zoomed, resulting in a more readable document.

There are several viewing products on the market today which use PDLs as the document data source, including Interleaf's WorldView product and Adobe's Acrobat. These products provide electronic documents that look just like their paper version but can also give users powerful tools for navigating through the documents to find the information they want. And because creating a PDL file usually is as easy as sending a document to a printer (in effect, you print to a file rather than to your hardcopy device), it is a very low cost solution. It is ideal for organizations that require that their original hardcopy document match their electronic pages and do not want to invest heavily in data preparation.



The standards-based approach can be the most powerful but also the most expensive.

There is an alternative PDL approach that does not simply put the hardcopy version of the document on screen but instead is optimized for screen delivery in terms of fonts, page size, formatting of structural elements. It may even include hypertext links or an entire "graphical navigation" system. (This is the approach Interleaf's WorldView takes.) Because this process of producing these enhanced PDL files can be highly automated, the production costs are about the same as for producing PDL files that replicate the paper versions, although there are some upfront design costs to design the on-line format and the navigation system. This approach gives users a familiar document "look and feel" to interact with while getting the benefit of having documents that are more readable, usable, and navigable on screen.

STANDARDS

The third approach—in addition to rasters and PDLs—is a standards-based approach, using SGML as the document data format. SGML enables data to be re-used across a variety of systems and easily maintained over a long lifecycle. Information used in these SGML-based retrieval systems can also be navigated through full text retrieval, boolean queries, hypertext, and other navigational aids as can PDL-based data.

The formatting information in an SGML-based system is interchanged using a FOSI (Formatting Output Specification Instance) by encoding it in SGML syntax. A distribution strategy using SGML should not require the replication of the original hardcopy. This is because the FOSI deals with a class of documents, like a DTD. This means that it is easy to convey, for example, that a set of documents should be in two-column format and that warnings should appear in bold. However, the FOSI is not designed, for instance, to specify where on the page a particular element was placed that a table broke to a new page at the fifteenth row.

This means that replicating the original hardcopy will be expensive because it will require manual intervention. CALS Output Specification, wherein the FOSI language is defined, is very clear on this point. It recommends that neither page fidelity (exact page matching) nor page integrity (matching the page breaks only), be an objective when interchanging CALS/1840A data.

Of all the various approaches, the standards-based approach can be the most powerful but also the most expensive because to become navigable and searchable (as in a typical IETM application), the SGML data needs to be carefully prepared, usually by highly-trained authors who, for example, insert the attributes they think future users will want to be able to search for. While the expense of preparing this information for on-line delivery is higher than the other two formats suggested this is the direction that many organizations are moving to because their investment in information is protected.

CHOOSING A PATH

Should the paper and electronic version be identical? That depends on the usual balance of benefits, needs and costs. If replicating the hardcopy is important, the best strategy will certainly be a raster or PDL-based systems. Where hardcopy matching is not a requirement, SGML-based systems can give tremendous flexibility and software independence.

At Interleaf, we recognize that customers needs will vary significantly even within one document set. Our objective is to give customers the full range of tools they need to support the distribution strategy that makes most sense for them. We offer our customers standards-based SGML tools for creating documents that can be used on any SGML-compliant distribution system. We offer a PDL-based distribution product, called WorldView, that is optimized for on-line distribution and can accept files from standard word processors such as MS Word and Word Perfect, SGML files, or PostScript files. This comprehensive strategy enables users to have the most cost effective and appropriate solution for their requirements. ■